

Does good memory help you win games?

James Burrige

Department of Mathematics, University of Portsmouth, Portsmouth, PO1 2UP, UK

Yu Gao and Yong Mao

School of Physics and Astronomy, University of Nottingham, Nottingham, NG7 2RD, UK

(Dated: May 12, 2015)

We present a simple game model where agents with different memory lengths compete for finite resources. We show by simulation and analytically that an instability exists at a critical memory length, and as a result, different memory lengths can compete and co-exist in a dynamical equilibrium. Our analytical formulation makes a connection to statistical urn models, and we show that temperature is mirrored by the agent's memory. Our analysis is easily generalisable to many other game models with implications that we briefly discuss.

PACS numbers: 02.50.Le, 89.75.Fb, 05.65.+b

Introduction – All successful forms of life must eventually engage in competition for resources. The equilibrium analysis of these competitions began with von Neumann [1] and Nash [2]. The theory of games has since found applications in genetics, ecology, economics and sociology [3–6]. Computational implementation of games leads to agent-based models, which may be of particular importance in understanding the behaviour of financial systems [6, 7]. For example, the particularly successful minority game model [8–11], captures the competition between intelligent agents with a restricted form of memory. Recent work suggest such games may be generalised leading to clearly separated regimes of behaviour [12]. In general, understanding the complex collective behaviour arising from the non-linear interactions between individuals is a major challenge for statistical physics [13, 14].

In this Letter we present a simple game model: in each round, individual agents pick one of two urns, each offering a stochastic yield to be shared by the pickers. An agent has a memory of these payouts for the previous τ rounds to aid its decision. Such stochastic yield sharing arises in animal foraging behaviour and stock trading [15], where the urns may represent different prey species, foraging patches, or stocks. Some form of intelligence is essential in order to compete [16]. As with the minority game [8], agents' memory in our model is a tool for decision making. However, unlike the minority game, memory in our model is used to make direct estimates of the highest paying choice, rather than to second guess opponents' next moves. In common with the thermal minority game [10], dynamical urn models [18], and some evolutionary games [19], our agents have a “temperature”, which captures the level of noise in the switches they make in search of yields. We find that the additional noise inherent in their finite memory samples, which is greater for shorter memories, leads the system to behave as if its agents have a higher temperature. Therefore, increasing memory “cools” the system. However, at a critical memory, a Hopf bifurcation [20] emerges producing stable cycles in the numbers of agents in each urn. Perhaps not surprisingly, a long memory is advan-

tageous, but the presence of these cycles allows short memory agents to compete, and a mixed memory system will evolve toward the bifurcation point. Our theoretical formulation follows that of statistical urn models such as Ehrenfest's dog flea model [21], which played an important role in the early development of statistical mechanics, and more recently allowed analytical investigation of effects such as slow relaxation and condensation in nonequilibrium statistical mechanics [18].

Whilst we have restricted this Letter to a yield sharing game, the analysis may be carried through for any social system where agents switch between strategies using their memory to determine the optimal choice. Simple examples include the Hawk Dove [3] and Rock Scissors Paper [22] games. Our formulation could also accommodate the inclusion of topological effects and agent interactions [23]. However, even without such complexities, urn models exhibit a remarkable range of non-equilibrium behavior that is connected with temperature [18]. Our analysis suggests that effects such as condensation and the emergence of order [24], which have social interpretation, may also have a connection with memory [23], but that long memory can also introduce instability.

Model definition – Consider the case of two urns and a total of n agents. We let the urn yields, $U_1(t), U_2(t)$, at round t be random variables uniformly distributed on $[0, \omega n]$ and $[0, n]$ respectively, where $\omega > 1$ so that urn 1 yields more on average than urn 2. We allow agents access to the arithmetic mean of the last τ payoffs, but we note that other forms of sampling could be used. Letting ϕ_t be the fraction of agents in urn 1, then the difference in the average payoffs between urn 1 and urn 2 is

$$\Delta_t := \frac{1}{\tau} \sum_{s=0}^{\tau-1} \left[\frac{U_2(t-s)}{n(1-\phi_{t-s})} - \frac{U_1(t-s)}{n\phi_{t-s}} \right]. \quad (1)$$

We refer to τ as the “memory length” of the agents. Agent dynamics is encoded in transition probabilities between urns, which are deterministic functions of Δ_t . At each round, each agent will switch urns using the proba-

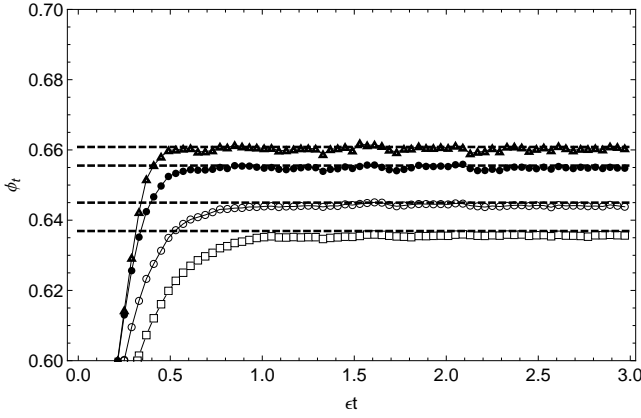


FIG. 1. Evolution of ϕ_t when $n = 10^6$, $\omega = 2$, $\beta = 10$ and $\epsilon = 10^{-6}$. Memory values are $\tau \in \{5, 10, 50, 500\}$ (squares, circles, dots, triangles). Dashed lines are analytical equilibrium values (see equation (12)).

bilities

$$W_{1 \rightarrow 2}(\Delta) = \frac{\epsilon}{2} [1 + \tanh(\beta\Delta)] \quad (2)$$

$$W_{2 \rightarrow 1}(\Delta) = \frac{\epsilon}{2} [1 - \tanh(\beta\Delta)]. \quad (3)$$

The parameter β , the “inverse temperature”, captures the degree of stochasticity with which agents make choices. For finite β , agents may decide to switch strategies even though their estimate of the payoff difference is unfavourable. In the limit $\beta \rightarrow \infty$ agents will only move if their estimate of the payoff difference indicates that the move is favourable. The parameter ϵ controls the rate at which strategy switching takes place compared to the rate at which yield information arrives. It may also be seen as the frequency with which opportunities to switch strategy arise. In the limit $\epsilon \rightarrow 0$, at most one agent will move at each round.

Simulation (instability) – We simulate the model for a series of values of τ when $n = 10^6$. Two different values of ϵ are used; in Figure 1 we have $\epsilon^{-1} = 10^6 \gg \tau$ and in Figure 2 we have $\epsilon = 10^{-3}$. For $\epsilon = 10^{-6}$, the expected number of moves at each step is < 1 , and ϕ appears very stable. For larger ϵ , ϕ experiences much larger fluctuations about the steady state value, driven by the yield process. For shorter memory values these fluctuations are random, but as τ approaches ϵ^{-1} , periodic oscillations appear and dominate. The appearance of these stable oscillations at critical memory, τ_c , is known as a Hopf bifurcation [20]. By allowing agents access only to the mean of their memory, we implicitly assume that changes in the expected payoff over the course of their memory, brought about by oscillations, are too subtle for them to infer from noise.

Simulation (coexistence) – We now investigate how agents with two different memories compete against one another by interpreting the payoff as reproduction rate. We define δ and γ as the rates of death, and reproduction

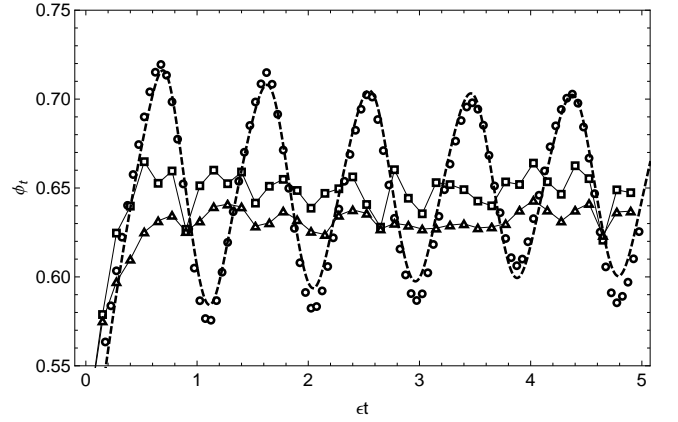


FIG. 2. Evolution of ϕ_t when $n = 10^6$, $\omega = 2$, $\beta = 5$ and $\epsilon = 10^{-3}$. Memory values are $\tau \in \{5, 50, 500\}$ (triangles, squares, circles). Dashed line is solution to equation (15) when $\tau = 500$ and ω, β, ϵ are as above.

per unit payoff, respectively. Reproduction is assumed to occur before death in each round, but in practice the probability of any one agent reproducing and dying in the same round is extremely small for the γ, δ values we choose. Letting $p_i^\tau(t)$ be the number of agents with memory τ in urn i at time t we set the probability of birth for an agent in urn i to be

$$\mathbf{P}(\text{birth}) = \frac{\gamma U_i(t)}{\sum_\tau p_i^\tau(t)}. \quad (4)$$

The death probability for each agent is set equal to δ . If populations are fixed in size and the system is not in an oscillatory state, then we expect that in equilibrium the longer memory agents will dominate the high yielding urn. Their long memory allows them to perceive smaller statistical advantages that are obscured by noise for the short memory agents. Using the thermodynamic analogy, the higher temperature (shorter memory) agents are more likely to make moves which leave them in an urn with a lower expected payoff, corresponding to a higher “energy” state. Above zero temperature, and in the absence of oscillations, the high yield urn will be under-exploited, placing high memory agents at an advantage. This effect can be observed in Figure 3 where we have simulated a mixed population of two memories $\tau \in \{10, 1000\}$ beginning with a ratio of 10:1 short memory to long memory agents. We see that initially the advantage afforded the long memory agents causes their population to grow, whereas the short memory agents reduce in number. Were this advantage to be sustained indefinitely then we would expect the short memory agents to eventually disappear, but in fact the populations stabilize. This effect appears because the long memory agents cause oscillations to develop once they are in sufficiently high concentration. In the presence of oscillations the short memory agents have an advantage because they can quickly observe opportunities offered by the oscillating

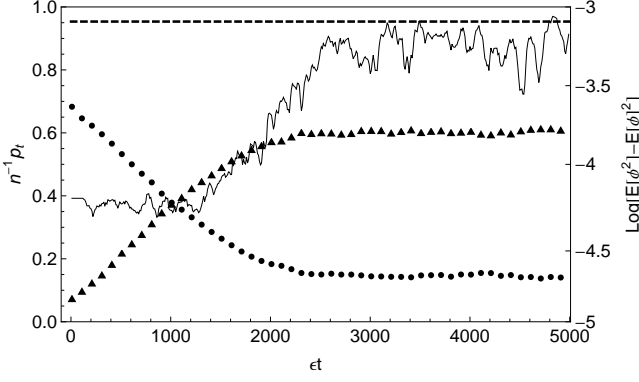


FIG. 3. Scaled populations $p(t) := p_1^\tau(t) + p_2^\tau(t)$ for $\tau = 10$ (circles) and $\tau = 1000$ (triangles) when $n = p(0) = 10^6$, $\epsilon = 10^{-3}$, $\beta = 4$ with $\gamma = 10^{-4}$ and $\delta = 2 \times 10^{-4}$. Also shown (thin black line) is evolution of variance of ϕ_t during population dynamics simulation. Straight dashed line shows variance of homogeneous population with the same ϵ, β, ω values at critical memory τ_c . Note: rapid initial equilibration of population values (bringing birth and death into balance) is not visible on time scale of plot.

payoffs. We therefore expect the system to evolve to the point where oscillations are just beginning to form. We may observe this evolution by making use of the variance of ϕ_t as an order parameter which captures proximity to the Hopf bifurcation point. In Figure 3 we see that at a critical ratio of short to long memory agents, the variance climbs rapidly, stabilizing just below the value seen in a system where all agents have memory τ_c but all other parameters are equal. In this way the Hopf bifurcation may be viewed as a self organized state.

Analysis (equilibrium) – We consider the behaviour of the model as $\epsilon \rightarrow 0$, allowing us to view it as an urn model in the Ehrenfest class [18] where agents independently make transitions using state (ϕ_t) dependent probabilities. Provided $\tau \ll \epsilon^{-1}$, the fraction ϕ_t may be approximated by a constant ϕ during the window over which payoff averaging takes place. In this case, by the central limit theorem, the marginal distributions of Δ_t for each t are approximately normal $N(\bar{\Delta}, \sigma^2/\tau)$ where, from (1)

$$\bar{\Delta}(\phi, \omega) := \frac{1}{2} \left(\frac{1}{1-\phi} - \frac{\omega}{\phi} \right) \quad (5)$$

$$\frac{\sigma^2(\phi, \omega)}{\tau} := \frac{1}{12\tau} \left[\frac{\omega^2}{\phi^2} + \frac{1}{(1-\phi)^2} \right]. \quad (6)$$

We now introduce an intermediate time scale T satisfying $\tau \ll T \ll \epsilon^{-1}$ and define the time average $\langle \cdot \rangle$, over a window of length T

$$\langle W_{i \rightarrow j}(\Delta) \rangle(t) := \frac{1}{T} \sum_{s=t-T+1}^t W_{i \rightarrow j}(\Delta_s). \quad (7)$$

This average is a random variable which, for constant ϕ , has expected value $\mathbf{E}[W_{i \rightarrow j}(\Delta)]$ where the expectation is

taken over the marginal distribution of Δ . The condition $\tau \ll T \ll \epsilon^{-1}$ ensures that ϕ is approximately constant over the window and that the variance of $\langle W_{i \rightarrow j}(\Delta) \rangle$ is proportional to T^{-1} (because Δ_{t_1} and Δ_{t_2} are dependent only when $|t_2 - t_1| < \tau \ll T$). As $\epsilon \rightarrow 0$, then assuming T is sufficiently large, the probability that an agent will make a transition $i \rightarrow j$ during interval T approaches $T \langle W_{i \rightarrow j}(\Delta) \rangle \approx T \mathbf{E}[W_{i \rightarrow j}(\Delta)]$, equivalent to a memoryless (Ehrenfest class) model where transition probabilities (2,3) are replaced with their expectations $\mathbf{E}[W_{i \rightarrow j}(\Delta)]$. Averaging over the normally distributed difference Δ we find that

$$\langle W_{1 \rightarrow 2}(\Delta) \rangle \approx \mathbf{E}[W_{1 \rightarrow 2}(\Delta)] \approx \frac{\epsilon}{2} [1 + \tanh(\alpha \bar{\Delta})] \quad (8)$$

where

$$\alpha = \sqrt{\frac{2\tau\beta^2}{2\tau + \pi\beta^2\sigma^2}}. \quad (9)$$

To obtain this result, we have made the approximation $\tanh(\beta\Delta) \approx \text{erf}(\sqrt{\pi}\beta\Delta/2)$, allowing us to make use of the exact relationship $\mathbf{E}[\text{erf}(\sqrt{\pi}\beta\Delta/2)] = \text{erf}(\sqrt{\pi}\alpha\bar{\Delta}/2)$. The constant α acts as an effective inverse temperature and we see that increasing τ “cools” the system closer to the inverse temperature β , and in the limit $\beta \rightarrow \infty$, $\alpha \propto \sqrt{\tau}$. To complete our analogy to a thermal urn model we now write the probability of finding the agents in a particular arrangement, or microstate, i , such that a fraction ϕ are in urn 1, as $p_i(\phi) \propto e^{-\alpha E}$ where E is an “energy” function. Considering two microstates separated by a single transition, and defining $\delta\phi = 1/n$, then detailed balance requires that in equilibrium $2\alpha\bar{\Delta} = \partial_\phi(\alpha E)\delta\phi$. This condition allows $E(\phi)$ to be computed, in principle, by integration. A closed form approximation $E(\phi) \approx -n \ln[\phi^\omega(1-\phi)]$ is obtained by noting that α depends weakly on ϕ compared to E so that $\partial_\phi(\alpha E) \approx \alpha \partial_\phi E$. Summing over all microstates corresponding to macrostate ϕ we have a Boltzmann probability distribution for ϕ

$$\mathbf{P}(\phi) = \frac{n!}{(n\phi)!(n(1-\phi))!} \frac{e^{-\alpha(\phi)E(\phi)}}{\mathcal{Z}}, \quad (10)$$

where \mathcal{Z} is the partition function. Taking the thermodynamic limit $n \rightarrow \infty$, and making use of Stirling’s approximation, we find that the most likely (maximum entropy) fraction, $\bar{\phi}$, satisfies:

$$\frac{1}{2n} \frac{\partial}{\partial \phi} \ln \mathbf{P}(\phi) = \alpha \bar{\Delta} - 2\phi + 1 = 0. \quad (11)$$

As the memory increases and the system cools we expect the agents to arrange themselves so that yields are shared more fairly. We therefore linearize (11) about the perfectly fair state, $\phi = \omega/(1+\omega)$, where agents in both urns receive the same expected payoff, finding that

$$\bar{\phi} \approx \frac{f(\tau) + \frac{\beta(1+\omega)^2}{2}}{2f(\tau) + \frac{\beta(1+\omega)^3}{2\omega}}, \quad (12)$$

where $f(\tau) = \sqrt{1 + \pi\beta^2(1 + \omega)^2/(12\tau)}$. The accuracy of this approximation is verified in Figure 1. For larger values of ϵ (Figure 2) agents move more quickly so the averaging effect (8) damps fluctuations in transition rates less strongly, creating larger fluctuations in ϕ_t . For finite β the system cannot reach perfect fairness for any memory length, but in the limit $\beta \rightarrow \infty$ where the transition probabilities (2,3) become step functions, we have that:

$$\bar{\phi} \approx \frac{\omega}{\omega + 1} \left[1 - \frac{\sqrt{\pi}(\omega - 1)}{\sqrt{3\tau}(\omega + 1)^2} + \mathcal{O}(\tau^{-1}) \right]. \quad (13)$$

From this we see that the distance away from the fair state decreases as $\tau^{-1/2}$ as the memory of the agents becomes large. However, we now show why increasing τ too far, when ϵ is finite, destabilizes the system.

Analysis (instability) – As τ increases, fluctuations in Δ_t due to the yield process are reduced but for finite ϵ we can no longer treat ϕ_t as a constant over the averaging window. It is instructive, therefore, to study the effect of variations in ϕ_t , neglecting the variations in yield. Promoting t to a continuous variable and replacing the urn yields with their mean values we have

$$\Delta_t \approx \frac{1}{2\tau} \int_{t-\tau}^t \left[\frac{1}{1 - \phi_s} - \frac{\omega}{\phi_s} \right] ds. \quad (14)$$

We then approximate the evolution of ϕ_t using the following delay differential equation:

$$\dot{\phi}_t = (1 - \phi_t)W_{2 \rightarrow 1}(\Delta_t) - \phi_t W_{1 \rightarrow 2}(\Delta_t). \quad (15)$$

A numerical solution to this equation is shown in Figure 2, along with simulation results using the same parameter values. The oscillations in the simulation are accurately captured by (15) but the stochastic yield disrupts their perfect periodicity. To discover the parameter values at which stable oscillations develop we linearize equation (15) by writing $\phi_t = \bar{\phi} + \psi_t$ where ψ_t are small fluctuations and $\bar{\phi}$ is the constant fixed point, not necessarily stable, of equation (15). In terms of these new variables

$$\Delta_t \approx \bar{\Delta}(\bar{\phi}, \omega) + 6 \frac{\sigma^2(\bar{\phi}, \sqrt{\omega})}{\tau} \int_{t-\tau}^t \psi_s ds \quad (16)$$

where the functions $\bar{\Delta}$ and σ^2 are defined in (5) and (6). After expanding the tanh functions in the transition rates to first order about $\bar{\Delta}(\bar{\phi}, \omega)$, we obtain the following linear delay equation

$$\dot{\psi}_t = -\epsilon \left[\psi_t + \frac{A}{\tau} \int_{t-\tau}^t \psi_s ds \right] \quad (17)$$

where $A = 3\beta \text{sech}^2[\beta \bar{\Delta}(\bar{\phi}, \omega)] \sigma^2(\bar{\phi}, \sqrt{\omega})$. To determine the stability of this equation we introduce an exponential trial solution $\psi_t = e^{\lambda t}$ where $\lambda = x + iy$. Substitution into equation (17) yields a characteristic equation with

real and imaginary parts given by

$$x^2 - y^2 + \epsilon x + \frac{\epsilon A}{\tau} (1 - e^{-\tau x} \cos \tau y) = 0 \quad (18)$$

$$2xy + \epsilon y + \frac{\epsilon A}{\tau} e^{-\tau x} \sin \tau y = 0. \quad (19)$$

For sufficiently small memory, τ , the real part, x , of the solutions to (18) and (19) is negative so the fixed point $\bar{\phi}$ is stable. As we increase τ , λ crosses through the imaginary axis, creating a switch to instability with oscillations of exponentially increasing magnitude. Although the full equation (15) shares this transition to instability, we find that the resulting oscillations are bounded. The appearance of these stable oscillations as τ passes through a critical value, which we denote τ_c , constitutes the Hopf Bifurcation [20]. To compute τ_c we set $x = 0$ in equation (19) so that $\text{sinc}(\tau y) = A^{-1}$. Expanding the sinc function to second order about its root at π/τ and solving the resulting quadratic we find that

$$y \approx \frac{\pi}{2\tau} \left(3 - \sqrt{1 - 4A^{-1}} \right) := \frac{\kappa}{\tau} \quad (20)$$

which defines a new constant κ . Substitution of this solution into (18), yields the following expression for the critical memory length

$$\tau_c = \frac{\kappa^2}{\epsilon A(1 - \cos \kappa)}. \quad (21)$$

For example, for the parameter values used in Figure 2, we have $\tau_c \approx 400$, whereas the relevant critical value for Figure 1 is $\tau_c = 1.8 \times 10^5$. These values are in excellent agreement with simulations.

Conclusion – We have introduced a simple thermal urn model of competition between agents with memory. Increasing memory allows agents to more accurately determine the most productive strategy, and reduces the temperature of the model. However, if a sufficiently high concentration of long memory agents is present a limit cycle appears which reduces the competitiveness of long memory agents, leading to self organized Hopf bifurcation in a mixed memory model. The simplicity of our model, its connection to classical urn models, together with the fact that limit cycles arise naturally suggest it might be fruitfully generalized, and employed to study a variety of different games. For example our approach may be applied to the Rock Scissors Paper game [3], where agents, interacting pairwise, recall their last τ interactions. Although a larger memory provides better statistical data on the optimal strategy, at critical memory a limit cycle emerges about Nash equilibrium, destroying this competitive advantage [27]. Other natural extensions include the introduction of multiple urns to represent, for example, different financial stocks. In this case we would expect more complex patterns of oscillation [26]. Experimental research into the nature of human and animal memory [28–31] places emphasis on the “forgetting function” which describes how memories decay with time. Such a function, or greater powers of statistical inference, could be naturally incorporated into our analysis, and its effects on stability explored.

-
- [1] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behaviour*, Princeton University Press (1953).
 - [2] J. Nash, *Annals of Mathematics*, **54**, 286-295 (1951).
 - [3] J. Maynard Smith, *Evolution and the theory of games*, Cambridge University Press (1982).
 - [4] R. B. Myerson, *Game Theory: Analysis of Conflict*, Harvard University Press (1991).
 - [5] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*, Cambridge University Press (1998).
 - [6] J. D. Farmer and D. Foley, *Nature*, **460**, 685-686 (2009).
 - [7] A. Chakraborti, D. Challet, A. Chatterjee, M. Marsili, Y.-C. Zhang and B. K. Chakrabarti, *Physics Reports*, **552** 1-25 (2015).
 - [8] D. Challet and Y.-C. Zhang, *Physica A*, **246**, 407-418 (1997).
 - [9] D. Challet, M. Marsili and Y. C. Zhang, *Minority Games*, Oxford University Press (2005).
 - [10] A. Cavagna, J. P. Garrahan and I. Giardina and D. Sherrington, *Phys. Rev. Lett.* **83** 4429 (1999).
 - [11] J. P. Garrahan, E. Moro and D. Sherrington, *Phys. Rev. E* **62** R9 (2000).
 - [12] T. Galla and J. D. Farmer, *PNAS*, **110**, 1232 (2013).
 - [13] J.-P. Bouchaud, *Physica A*, **313**, 238-251 (2002).
 - [14] R. Mantegna and H. E. Stanley, *An Introduction to Econophysics*, Cambridge University Press (1999).
 - [15] S. Saavedra, R. D. Mailmgen, N. Switanek and B. Uzzi, *P. R. Soc. B.* **280** 20122901 (2013).
 - [16] C. Belisle and J. Cresswell, *Theor. Popul. Biol.* **52** 78-90 (1997).
 - [17] A. Cavagna, *Phys. Rev. E.*, **59** 3783 (1999).
 - [18] C. Godreche and J. M. Luck, *J. Phys. Cond. Matt.* **14** 1601-1615 (2002).
 - [19] A. Traulsen, J. M. Pacheco and M. A. Nowak, *J. Theor. Biol.* **246** 522-529 (2007).
 - [20] T. Erneux, *Applied Delay Differential Equations*, Springer (2009).
 - [21] P. Ehrenfest, *Phys. Zeit.*, **8**, 311 (1907).
 - [22] Z. Wang, B. Xu and H.J. Zhou, *Scientific Reports* **4** 5830 (2014).
 - [23] F. Caccioli, S. Franz and M. Marsili, *J. Stat. Mech: Theory and Experiment*, **7**, P07006 (2008).
 - [24] C. Castellano, S. Fortunato and V. Loreto, *Rev. Mod. Phys.* **81** (2009).
 - [25] H. Mahmoud, *Polya urn Modles*, CRC Press, Boca Raton (2008).
 - [26] P. Manfredi and L. Fanti, *Economic Modelling*, **21** 573-594 (2004).
 - [27] J. Burridge, In Preparation.
 - [28] W. K. Estes, *J. Am. Stat. Assoc.*, **67** 81-102 (1972).
 - [29] P. Killeen, *Psychon. B. Rev.* **8** 18-43 (2001).
 - [30] R. G. Cook and A. P. Blaisdell, *Behav. Proc.* **72** 255-264 (2006).
 - [31] L. Averell and A. Heathcote, *J. Math. Psychol.* **55** 25-35 (2011).